

CLAIMS

- 1 1. An extensible, system-independent, version-interoperable format for transmitting
2 a data stream having data set information from a source system to a replica residing on a
3 destination system comprising:
4 a plurality of standalone headers having discrete identifiers, each of the plurality
5 of standalone headers being representative of a plurality of data stream characteristics;
6 and
7 a data following header that follows, in the data stream, the plurality of standalone
8 headers and that indicates that the data set information is following the data following
9 header, the data set header including an extended attribute field that associates an ex-
10 tended attribute with the data set information.
- 1 2. The format as set forth in claim 1 wherein the plurality of standalone headers each
2 include an indication of one of a plurality of specialized header types and at least some of
3 the plurality of specialized header types are adapted for carrying directory inode data.
- 1 3. The format as set forth in claim 3 wherein the data stream is adapted to carry
2 source file system inode data and source file generation numbers.
- 1 4. The format as set forth in claim 2 wherein one of the specialized header types
2 comprises a deleted files type and the directory inode data comprises a list of deleted files
3 on the source file system.
- 1 5. The format as set forth in claim 1 wherein the extended attributes include ACLs
2 and streams associated with a plurality of operating systems and system architectures.
- 1 6. The format as set forth in claim 1 wherein one of the plurality of standalone head-
2 ers comprises an open file/undo header that instructs the destination system to revert to an
3 earlier copy of a stored file identified by the open file/undo header.

1 7. The format as set forth in claim 1 wherein the data set information comprises file
2 information.

1 8. The format as set forth in claim 1 wherein the data set information comprises
2 changed files on the source system transmitted for backup on the replica of the destina-
3 tion system.

1 9. The format as set forth in claim 1 wherein the data following header includes off-
2 set and block number information with respect to the data set information that follows the
3 data following header.

1 10. The format as set forth in claim 1 wherein data following header comprises a
2 fixed-length record including (a) a generic part for storing an indication of a data follow-
3 ing header type; (b) a non-generic part, adapted to carry predetermined data related to the
4 extended attribute and data related to offsets and block numbers for the data set informa-
5 tion that follows the data following header; and (c) a space for a bit-code representative
6 of a name associated with the extended attribute.

1 11. The format as set forth in claim 1 wherein each of the plurality of standalone
2 headers comprises a fixed-length record including a generic part for storing an indication
3 of one of a plurality of specialized header types, a non-generic part, adapted to carry pre-
4 determined data related one of the specialized header types and a space for additional in-
5 formation.

1 12. The format as set forth in claim 1 wherein the data following header is adapted to
2 be positioned within the data stream at predetermined intervals that are up to approxi-
3 mately 2 MB of data set information in size.

1 13. The format as set forth in claim 1 wherein the destination system is adapted to
2 receive the data following header with the extended attribute and cause the data set in-
3 formation associated with the extended attribute to be stored an entry in a hidden perma-
4 nent metadirectory with identifiers that are the same as identifiers for the data set infor-
5 mation in a file system of the destination system, the entry having the extended attribute
6 associated therewith so that retrieval of the entry from the hidden permanent metadirec-
7 tory also retrieves the extended attribute.

1 14. The format as set forth in claim 13 wherein the destination system also includes a
2 hidden purgatory metadirectory in which current data set information from the hidden
3 permanent directory is stored during an update of the hidden permanent metadirectory
4 with changed data set information, the destination system being further adapted to (a)
5 delete the hidden purgatory metadirectory after a complete receipt of all expected
6 changed data set information of the hidden permanent metadirectory with the changed
7 data set information, and (b) move current data set information stored on the hidden pur-
8 gatory directory back to the hidden permanent metadirectory after an incomplete receipt
9 of all expected changed data set information.

1 15. The format as set forth in claim 14 wherein the destination system is adapted to
2 create hidden new metadirectory to store changed data set information for transfer to the
3 hidden permanent directory after of the complete receipt of all the expected changed data
4 set information.

1 16. The format as set forth in claim 1 wherein the source system and the destination
2 system are remote with respect to each other and interconnected by a network, and
3 wherein the data stream is encapsulated within a networking protocol adapted for trans-
4 mission over the network.

1 17. A format for transmitting a data stream that includes data set information between
2 a source system and a replica stored on the destination system comprising:

3 a data following header appended to a predetermined-sized chunk of data, the data
4 following header including a field that identifies extended attributes associated with data
5 set information carried in the chunk.

1 18. The format as set forth in claim 17 wherein the extended attributes include ACLs
2 and streams associated with a plurality of operating systems and system architectures.

1 19. The format as set forth in claim 17 wherein the data set information comprises file
2 information.

1 20. The format as set forth in claim 17 wherein the data set information comprises
2 changed files on the source system transmitted for backup on the replica of the destina-
3 tion system.

1 21. The format as set forth in claim 17 wherein the data following header includes
2 offset and block number information with respect to the data set information that follows
3 the data following header.

1 22. The format as set forth in claim 17 wherein data following header comprises a
2 fixed-length record including (a) a generic part for storing an indication of a data follow-
3 ing header type; (b) a non-generic part, adapted to carry predetermined data related to the
4 extended attribute and data related to offsets and block numbers for the data set informa-
5 tion that follows the data following header; and (c) a space for a bit-code representative
6 of a name associated with the extended attribute.

1 23. The format as set forth in claim 17 wherein the chunk has a size of up to ap-
2 proximately 2 MB of data set information.

1 24. The format as set forth in claim 17 wherein the destination system is adapted to
2 receive the data following header with the extended attribute and cause the data set in-

3 formation associated with the extended attribute to be stored an entry in a hidden perma-
4 nent metadirectory with identifiers that are the same as identifiers for the data set infor-
5 mation in a file system of the destination system, the entry having the extended attribute
6 associated therewith so that retrieval of the entry from the hidden permanent metadirec-
7 tory also retrieves the extended attribute.

1 25. The format as set forth in claim 24 wherein the destination system also includes a
2 hidden purgatory metadirectory in which current data set information from the hidden
3 permanent directory is stored during an update of the hidden permanent metadirectory
4 with changed data set information, the destination system being further adapted to (a)
5 delete the hidden purgatory metadirectory after a complete receipt of all expected
6 changed data set information of the hidden permanent metadirectory with the changed
7 data set information, and (b) move current data set information stored on the hidden pur-
8 gatory directory back to the hidden permanent metadirectory after an incomplete receipt
9 of all expected changed data set information.

1 26. The format as set forth in claim 25 wherein the destination system is adapted to
2 create hidden new metadirectory to store changed data set information for transfer to the
3 hidden permanent directory after of the complete receipt of all the expected changed data
4 set information.

1 27. A method for storing and retrieving extended attributes associated with a data set
2 information comprising:

3 storing a current data set information with current extended attributes in a perma-
4 nent hidden metadirectory;

5 transferring the data set information to a purgatory metadirectory upon receipt of
6 a changed data set information;

7 storing the received changed data set information in a new metadirectory; and

8 upon completion of receipt of all expected changed data set information, transfer-
9 ring the received changed data set information from the new metadirectory to the perma-

10 nent metadirectory, the permanent metadirectory thereby being available for retrieval of
11 extended attributes associated with the data set information.

1 28. The method as set forth in claim 27 further comprising the step of, upon a failure
2 to complete receipt of all expected changed data set information, transferring the current
3 data set information with the current extended attributes back to the permanent metadi-
4 rectory.

1 29. The method as set forth in claim 28 wherein the data set information comprises
2 files organized in a directory tree structure the same as a file system structure on the des-
3 tination system and wherein the extended attributes comprise ACLs and streams associ-
4 ated with the files.

1 30. The method as set forth in claim 29 further comprising the step of deleting the
2 purgatory metadirectory after one of, either (a) the transferring of the changed data set
3 information from the new metadirectory to the permanent metadirectory or (b) the trans-
4 ferring of the current data set information from the purgatory metadirectory back to the
5 permanent metadirectory.

1 31. The method as set forth in claim 30 further comprising, upon a request from the
2 source to restore data sets from the data set information, scanning the permanent direc-
3 tory and retrieving the data sets including retrieving respective of the extended attributes
4 associated with the data sets.

1 32. The method as set forth in claim 31 further comprising the step of providing the
2 retrieved data sets' extended attributes in a format for transmission to the source from the
3 destination, the format including data following headers each having a field that associ-
4 ates the respective of the extended attributes with the retrieved data sets.

1 33. The method as set forth in claim 32 wherein the respective of the extended attrib-
2 utes are associated with the data sets based upon NT streams.

1 34. The method as set forth in claim 27 wherein the extended attributes are associated
2 with the data set information in the permanent metadirectory using NT streams.

1 35. The method as set forth in claim 27 further comprising the step of providing the
2 data sets' extended attributes in a format for transmission to the destination from the
3 source, the format including data following headers each having a field that associates the
4 respective of the extended attributes with the retrieved data sets.

1 36. A method for ensuring coherency in a data set transmitted from a source system to
2 a replica on a destination system comprising the steps of:
3 retrieving a first modification time on the source system for the data set;
4 opening the data set on the source system and transmitting the data set from the
5 source system to the destination;
6 after completing transmitting, closing the data set on the source system and re-
7 trieving second modification time on the source system; and
8 if the second modification time and the first modification time are not the same,
9 providing by the source system an instruction to the destination system to revert to an
10 earlier stored copy of the data set on the replica.

1 37. The method as set forth in claim 36 wherein the step of providing the instruction
2 comprises transmitting an undo standalone header in a data stream that includes the data
3 set, the standalone header identifying the data set and indicating an undo header type.